SIDARTHa
European Emergency Data-based Syndromic Surveillance System

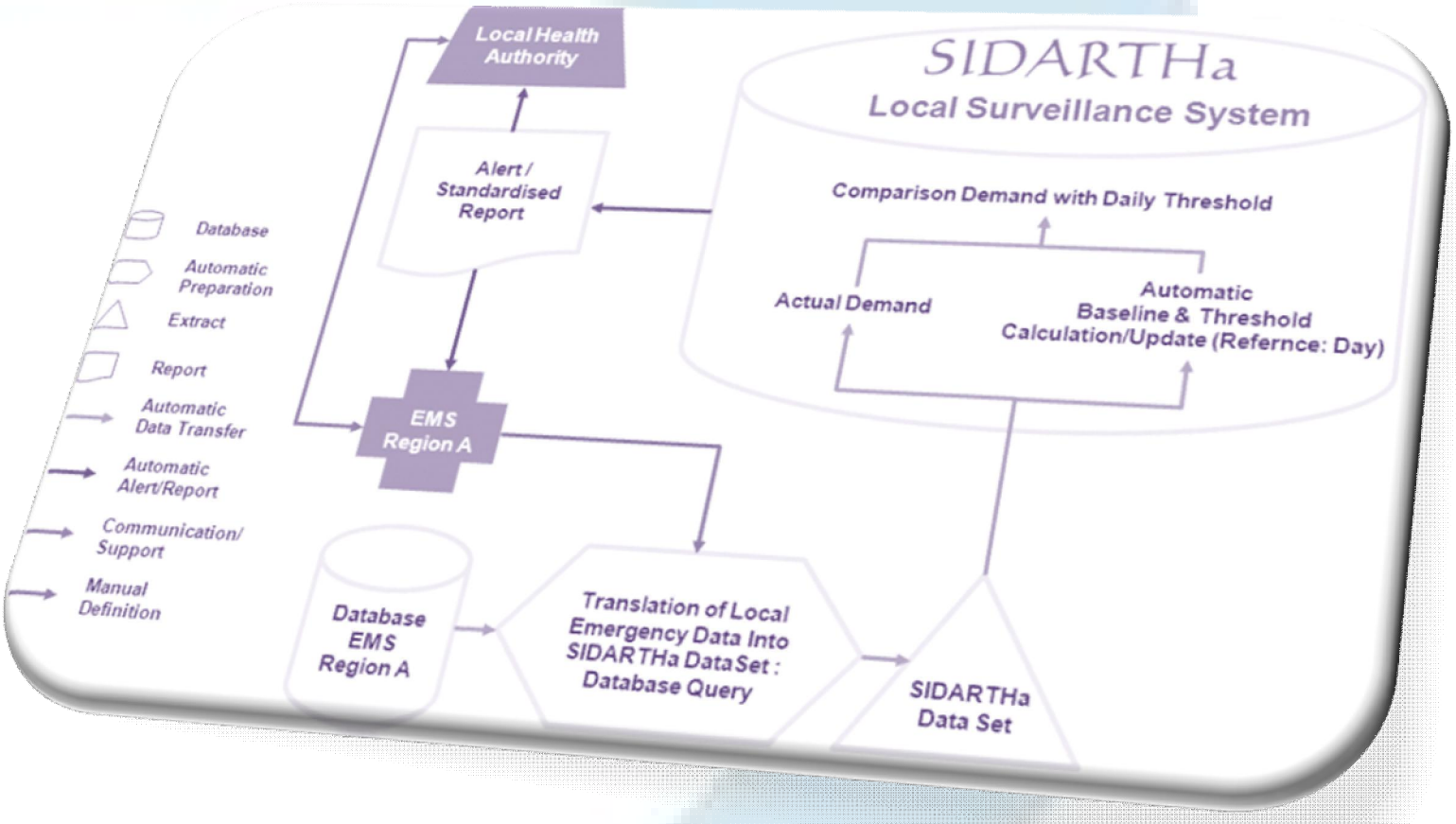
Executive Agency for Health and Consumers

Directorate-General for Health & Consumers

SIDARTHa
European Emergency Data-based Syndromic Surveillance System
Grant Agreement No. 2007208

SIDARTHa Technical Guidelines

How to use the European Emergency Data-based Syndromic Surveillance System SIDARTHa



SIDARTHa
European Emergency Data-based Syndromic Surveillance System

SIDARTHa - European Emergency Data-based Syndromic Surveillance System

The project 'European Emergency Data-based System for Information on, Analysis and Detection of Risks and Threats to Health – SIDARTHa' is co-funded by the European Commission under the Programme of Community Action in the Field of Public Health 2003-2008 (Grant Agreement-No.: 2007208).

SIDARTHa Steering Committee

Luis Garcia-Castrillo Riesgo (Project Leader), Thomas Krafft (Scientific-Technical Coordinator), Matthias Fischer, Alexander Krämer, Freddy Lippert, Gernot Vergeiner

SIDARTHa Project Group

Dispatch Centre Tyrol (Austria), contact person: Gernot Vergeiner; Federal Government, Department of Public Health (Belgium), contact person: Agnes Meulemans; Emergency Medical Service Prague (Czech Republic), contact person: Milana Pokorna; Capital Region (Denmark), contact person: Freddy Lippert; University Hospital Kuopio (Finland), contact person: Jouni Kurola; Emergency Medical Service Province Hauts de Seine (France), contact person: Michel Baer; Hospitals of County of Goeppingen (Germany), contact person: Matthias Fischer; GEOMED Research Forschungsgesellschaft mbH (Germany), contact person: Thomas Krafft; University of Bielefeld, Department of Public Health Medicine (Germany), contact person: Alexander Krämer; National Emergency Medical Service (Hungary), contact person: Gabor Göbl; Maastricht University, Department of International Health, contact person: Helmut Brand (Netherlands); San Martino University Hospital Genoa (Italy), contact person: Francesco Bermano; Haukeland University Hospital Bergen (Norway), contact person: Guttorm Brattebo; University of Cantabria (Spain), contact person: Luis Garcia-Castrillo Riesgo; University Hospital Antalya (Turkey), contact person: Hakan Yaman

Advisory Board

Helmut Brand (The Netherlands, Chair), Andrea Ammon (ECDC), Enrico Davoli (WHO-Euro), Per Kulling (2008-2009) & Franz Karcher (2009-2010) (EU Health Threat - Unit), Javier Llorca (Spain), Jerry Overton (USA), Santiago Rodriguez (Spain), Mark Rosenberg (Canada)

Coordination Office

Alexandra Ziemann (Science Officer), Weyma Notel (Project Assistant), Juan-José San Miguel Roncero (Financial Officer)

SIDARTHa Technical Guidelines
How to use the European Emergency Data-based Syndromic Surveillance System SIDARTHa

This report describes the technical setup and provides a guideline for implementation of the SIDARTHa syndromic surveillance system for early public health threat detection and risk communication and forms deliverable D7 as defined in the Grant Agreement.

Compiled by

Luis Garcia-Castrillo Riesgo, Alexandra Ziemann and Ronny Conde

Editors

Luis Garcia-Castrillo Riesgo, Alexandra Ziemann, Thomas Krafft, Ronny Conde, Matthias Fischer, Gernot Vergeiner, Freddy Lippert, Alexander Krämer, Helmut Brand for the SIDARTHa project group

Please cite as:

Garcia-Castrillo Riesgo L, Ziemann A, Krafft T, Conde R, Fischer M, Vergeiner G, Lippert F, Krämer A, Pinheiro P, Brand H for the SIDARTHa project group (ed.) (2010): SIDARTHa Technical Guidelines. How to implement the European Emergency Data-based Syndromic Surveillance System SIDARTHa. Bad Honnef.

Cover Figure

SIDARTHa Syndromic Surveillance System Design (own creation)

© SIDARTHa 2010

SIDARTHa Scientific/Technical Coordination Office, c/o GEOMED Research Forschungsgesellschaft mbH, Hauptstr. 68, 53604 Bad Honnef, Germany, Tel. +49 2224 7799896, Fax. +49 2224 7799897, mail@sidartha.eu, www.sidartha.eu

Contents

CONTENTS III
FIGURES IV
ABBREVIATIONS V
ACKNOWLEDGEMENTS VII

1 INTRODUCTION: THE SIDARTHA PROJECT 1

2 OBJECTIVES & METHODOLOGY 3

2.1 OBJECTIVES 3
2.2 METHODOLOGY 3

3 SIDARTHA SYNDROMIC SURVEILLANCE SYSTEM DESIGN 5

3.1 STATE-OF-THE-ART SYNDROMIC SURVEILLANCE SYSTEM DESIGN 5
3.2 SIDARTHA SYNDROMIC SUR-VEILLANCE SYSTEM DESIGN 6
3.3 SIDARTHA SYNDROMIC SUR-VEILLANCE APPLICATION 8
3.4 FUTURE USER REQUIREMENTS 9

4 TECHNICAL GUIDELINES 10

4.1 DEMO VERSION & SUPPORT 10
4.2 DOWNLOAD & SETUP 11
4.2 USER MANUAL 13
4.2.1 SIDARTHA LOCAL HEALTH SURVEILLANCE SYSTEM (LHSS) 13
4.2.2 SIDARTHA EUROPEAN HEALTH SURVEILLANCE SYSTEM (EHSS) 18

REFERENCES 19

APPENDIX 20

APPENDIX 1 SIDARTHA SHORT SURVEY ON USER REQUIREMENTS 21
APPENDIX 2 TECHNICAL DEVELOPMENT OF BEVALLEY HEALTH 25
APPENDIX 3 SIDARTHA APPLICATION SETUP: DATA BASE CONTENTS 32

© SIDARTHa 2010

Figures

FIGURE 1: SIDARTHA PROJECT METHODOLOGY 2
FIGURE 2: SIDARTHA APPROACH 2
FIGURE 3: COMPONENTS OF SYNDROMIC SURVEILLANCE SYSTEMS 5
FIGURE 4: SIDARTHA SYNDROMIC SURVEILLANCE SYSTEM DESIGN (VERSION 2) 7
FIGURE 5: SIDARTHA SYNDROMIC SURVEILLANCE SYSTEM DESIGN – REGIONAL LEVEL (VERSION 2) 7
FIGURE 6: TECHNICAL SETUP OF THE SIDARTHA SYNDROMIC SURVEILLANCE APPLICATION 8
FIGURE 7: SIDARTHA SHORT SURVEY RESULTS 9
FIGURE 8: SIDARTHA APPLICATION: FILE FORMAT 12

© SIDARTHa 2010

Abbreviations

14 - AKFAM – TR	SIDARTHa associated partner abbreviation for the University Hospital Antalya, Turkey
ARIMA	Autoregressive integrated moving average
Bvh	BeValley Health
C 1, C2, C3	Detection Algorithm used in the Early Aberration Reporting System
CUSUM	Cumulative Sum
D7	Deliverable No. 6 of the SIDARTHa project
dWh	Data Ware House
ECDC	European Centre for Disease Prevention and Control
ED	Emergency Department
EED	European Emergency Data
EHSS	European Health Surveillance System (in SIDARTHa)
EMD	Emergency Medical Dispatch
EMS	Emergency Medical Service
ETL	Extract-Transfer-Load
EU	European Union
FAQ	Frequently Asked Questions
8 – FOD Health DG 1 – BE	SIDARTHa associated partner abbreviation for the federal public service health, food chain safety and environment, Belgium
2 - GEOMED – DE	SIDARTHa associated partner abbreviation for the GEOMED Research Forschungsgesellschaft mbH, Germany
GIS	Geographic Information System
16 – HSanMartino – IT	SIDARTHa associated partner abbreviation for the San Martino University Hospital Genoa, Italy
17 – HUS – NO	SIDARTHa associated partner abbreviation for the Haukeland University Hospital Bergen, Norway
ICD	International Classification of Diseases
ILI	Influenza-Like-Illness
IT	Information Technology
7 – KAE – DE	SIDARTHa associated partner abbreviation for the Klinik am Eichert (Clinics of the County of Goeppingen), Germany
9 – KUH – FI	SIDARTHa associated partner abbreviation for the University Hospital Kuopio, Finland
LHSS	Local Health Surveillance System (in SIDARTHa)
4 – LST Tirol – AU	SIDARTHa associated partner abbreviation for the Dispatch Centre Tyrol, Austria
M	Month of the SIDARTHa project
MedISys	Medical Intelligence System
NLP	Natural Language Processing
5 – OMSZ – HU	SIDARTHa associated partner abbreviation for the National Emergency Medical Service Hungary
3 – RegH – DK	SIDARTHa associated partner abbreviation for the Capital Region Denmark
6 – SAMU – FR	SIDARTHa associated partner abbreviation for the System of Emergency Medical Assistance Garches, France
SIDARTHa	European Emergency Data-based System for Information on, Detection and Analysis of Risks and Threats to Health
SOA	Service Oriented Architecture
SQL	Structured Query Language

© SIDARTHa 2010

15 – UNIBI – DE	SIDARTHa associated partner abbreviation for the University of Bielefeld, Germany
1 – UNICAN – ES	SIDARTHa associated partner abbreviation for the University of Cantabria, Spain
URL	Uniform Resource Locator
USA	United States of America
WP	Work Package of the SIDARTHa project
WHO	World Health Organization
13 – ZZSHMP – USZS – CZ	SIDARTHa associated partner abbreviation for the Emergency Medical Service Prague, Czech Republic

© SIDARTHa 2010

Acknowledgements

This report presents the results of the Work Package (WP) 6 of the European Commission co-funded project "SIDARTHa – European Emergency Data-based Syndromic Surveillance System" and forms deliverable D7 as defined in the Grant Agreement (No. 2007208). The results presented in this report were compiled by 1 – UNICAN – ES, *BeValley Technologies*, and the Scientific/Technical Coordination Office. The authors would like to thank all project group members for their contributions:

Country Consortia (leading organisations/Associated Partners)

- **Austria:** Dispatch Centre Tyrol, Ing. Gernot Vergeiner, Andreas Maurer (4 – ILL GmbH – AU)
- **Belgium:** Federal Government, Department of Public Health, Dr. Agnes Meulemans, Dr. Jean Bernard Gillet (8 – FOD Health DG 1 – BE)
- **Czech Republic:** Emergency Medical Service Prague, Dr. Milana Pokorná, Dr. Petr Zajíček (13 – ZZSHMP-USZS – CZ)
- **Denmark:** Capital Region, Prof. Freddy Lippert (3 – RegH – DK)
- **Finland:** University Hospital Kuopio, Dr. Jouni Kurola, Dr. Tapio Kettunen (9 – KUH – FI)
- **France:** Emergency Medical Service Province Hauts de Seine, Dr. Michel Baer, Dr. Anna Ozguler (6 – SAMU – FR)
- **Germany:** Hospitals of the County of Goeppingen, Prof. Matthias Fischer, Dr. Martin Messelken (7 – KAE – DE)
- **Hungary:** National Emergency Medical Service, Dr. Gábor Gőbl (5 – OMSZ – HU)
- **Italy:** San Martino University Hospital Genoa, Prof. Francesco Bermano, Dr. Lorenzo Borgo (16 – HSanMartino – IT)
- **Norway:** Haukeland University Hospital Bergen, Dr. Guttorm Brattebø, Lars Myrmel (17 – HUS – NO)
- **Spain:** University of Cantabria, Prof. Luis Garcia-Castrillo Riesgo, Weyma Notel, Juan José San Miguel Roncero, Prof. Francisco Javier Llorca Diaz (1 – UNICAN – ES)
- **Turkey:** University Hospital Antalya, Dr. Hakan Yaman, Sercan Bulut (14 - AKFAM – TR)

The country consortia consist of emergency medical care institutions and local/regional public health authorities.

New Country Consortium (leading organisation, from September 2009)

- Netherlands: Maastricht University, Department of International Health, Prof. Helmut Brand, Dr. Genc Burazeri, Nicole Rosenkötter, Janneke Kraan

Scientific-Technical partners (Associated Partners/Technical Unit)

- GEOMED Research Forschungsgesellschaft mbH (Germany): Dr. Thomas Krafft, Alexandra Ziemann, Tim Tenelsen, Martina Schorbahn, Nico Reinke, Dr. Axel Kortevoß (2 – GEOMED – DE)
- University of Bielefeld, Department of Public Health Medicine (Germany): Prof. Alexander Krämer, Dr. Paulo Pinheiro (15 – UniBi – DE)
- Netherlands: Maastricht University, Department of International Health, Prof. Helmut Brand, Dr. Genc Burazeri, Nicole Rosenkötter, Janneke Kraan

External Scientific Advisory Board

- Prof. Helmut Brand, Maastricht University (Netherlands, Chair)
- Dr. Andrea Ammon, European Centre for Disease Prevention and Control (Sweden)
- Dr. Enrico Davoli, World Health Organization Regional Office for Europe, Division of Country Health Systems (Spain)
- Dr. Per Kulling/Dr. Franz Karcher, European Commission, Health Threat Unit (Luxembourg)

© SIDARTHa 2010

- Prof. Francisco Javier Llorca Diaz, University of Cantabria (Spain)
- Jerry Overton, MPA, International Academies of Emergency Dispatch (USA)
- Dr. Santiago Rodriguez, Health Service Cantabria (Spain)
- Prof. Mark Rosenberg, Queen's University, Department of Geography and Department of Community Health and Epidemiology (Canada)

© SIDARTHa 2010

1 Introduction: The SIDARTHa Project

Syndromic surveillance can detect public health threats earlier than traditional surveillance and reporting systems. Pre-hospital emergency medical services (EMS) and emergency medical dispatch centres (EMD), and in-hospital emergency departments (ED) across Europe routinely collect electronic data that provides the opportunity to be used for near real time syndromic surveillance of communicable and non-communicable health threats such as heat-related diseases or Influenza-Like-Illness (ILI). The European Commission co-funded project SIDARTHa (Grant Agreement No. 2007208) for the first time systematically explores the use of emergency data to provide a basis for syndromic surveillance in Europe. The project runs from June 2008 until December 2010. It is an initiative of emergency medical professionals organised in the *European Emergency Data (EED) – Research Network*[1].

Objectives

The objective of the European project SIDARTHa is to conceptualise, develop, implement/test and evaluate the European Emergency Data-based System for Information on, Detection and Analysis of Risks and Threats to Health (SIDARTHa).

Methodology

During the conceptualisation phase, information on international state-of-the-art in the early detection of health threats and on the current practice of health surveillance and alert systems in Europe are brought together with the possibilities of emergency data for detection of health threats and specific public health authority and emergency professional desires for SIDARTHa's system features. On this basis the surveillance system SIDARTHa is tested and evaluated during the implementation phase in different regions[2] (cf. Figure 1).

The project group constitutes a high-level expert panel of emergency professionals, public health experts and health authority representatives under guidance of an interdisciplinary steering committee. A sequence of focused methods such as group discussions, Strengths - Weaknesses - Opportunities - Threats analysis of existing procedures, half-standardised surveys to seek input from potential futures users, statistical analyses and modelling, and geo-processing methods are applied.

Expected Results & Products

The SIDARTHa project provides a methodology and software application for syndromic surveillance at the regional level[3] in Europe based on routinely collected emergency data. The SIDARTHa syndromic surveillance system automatically analyses the actual demand for emergency services and detects temporal and spatial aberrations from the expected demand. The system will automatically alert decision makers in the emergency medical institution and the regional public health authority. Via the established reporting ways the regional public health authority can inform national or supranational authorities on an event (cf. Figure 2).

It is expected that SIDARTHa improves the timeliness and cost-effectiveness of European and national health surveillance by providing a basis for systematic syndromic surveillance that supplements the existing surveillance structures.

The main outputs of the project are a syndromic surveillance application (software) publicly available free-of-charge and guidelines for future users on how to use the application and how to transform emergency data into syndromes and into the common SIDARTHa data set that the application can analyse, including recommendations on technical infrastructure, reporting procedures and interpretation of the results. Furthermore, the guidelines cover the utilisation of the interactive user display and risk communication platform.

[1] www.eed-network.eu

[2] SIDARTHa Implementation sites: District of Kufstein, Austria; Capital Region, Denmark, County of Goeppingen, Germany, Autonomous Region Cantabria, Spain

[3] In the SIDARTHa project the term *regional* is used referring to the smallest administrative level at which a health authority responsible for surveillance and reporting is established in a European country depending on the national definition and rules. This level can be a community, city, county, district or state. The implementation of the SIDARTHa syndromic surveillance system can be based on data collected for the same administrative level or also for a part of this area or based on the catchment areas of one or more participating emergency institutions.

© SIDARTHa 2010

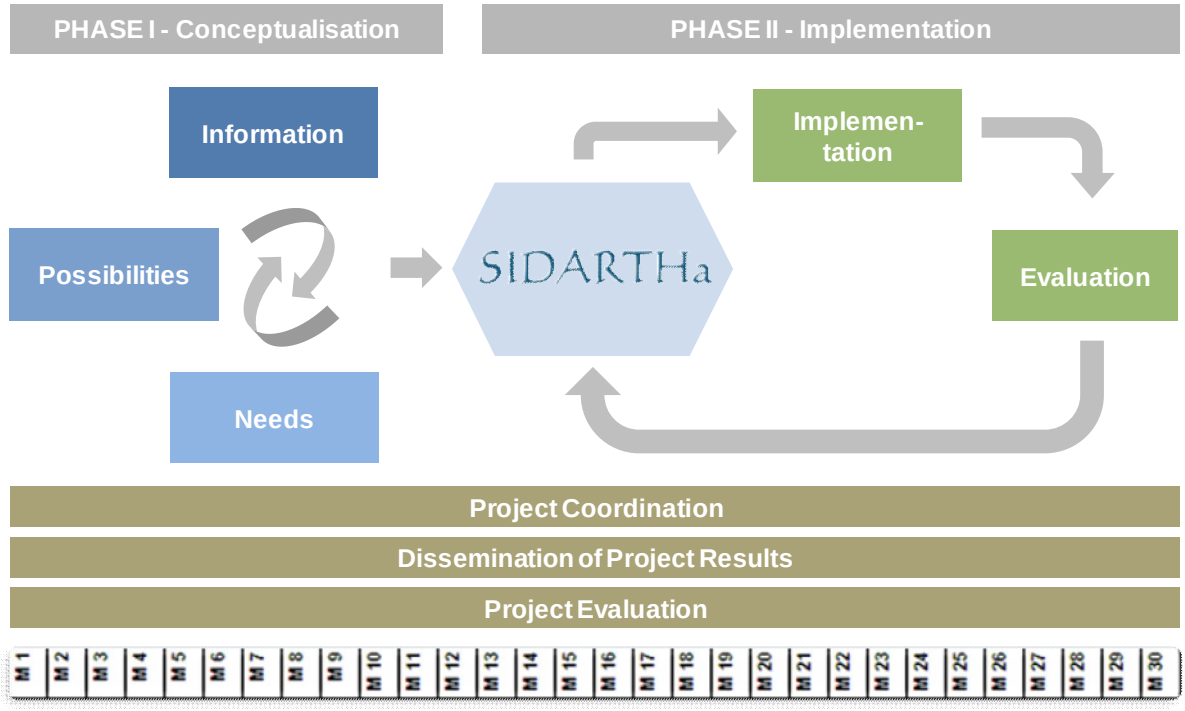

PHASE I - Conceptualisation
PHASE II - Implementation
Information
Possibilities
Needs
SIDARTHa
Implemen-tation
Evaluation
Project Coordination
Dissemination of Project Results
Project Evaluation
M 1 M 2 M 3 M 4 M 5 M 6 M 7 M 8 M 9 M 10 M 11 M 12 M 13 M 14 M 15 M 16 M 17 M 18 M 19 M 20 M 21 M 22 M 23 M 24 M 25 M 26 M 27 M 28 M 29 M 30

Figure 1: SIDARTHa Project Methodology

M = Month of the project time

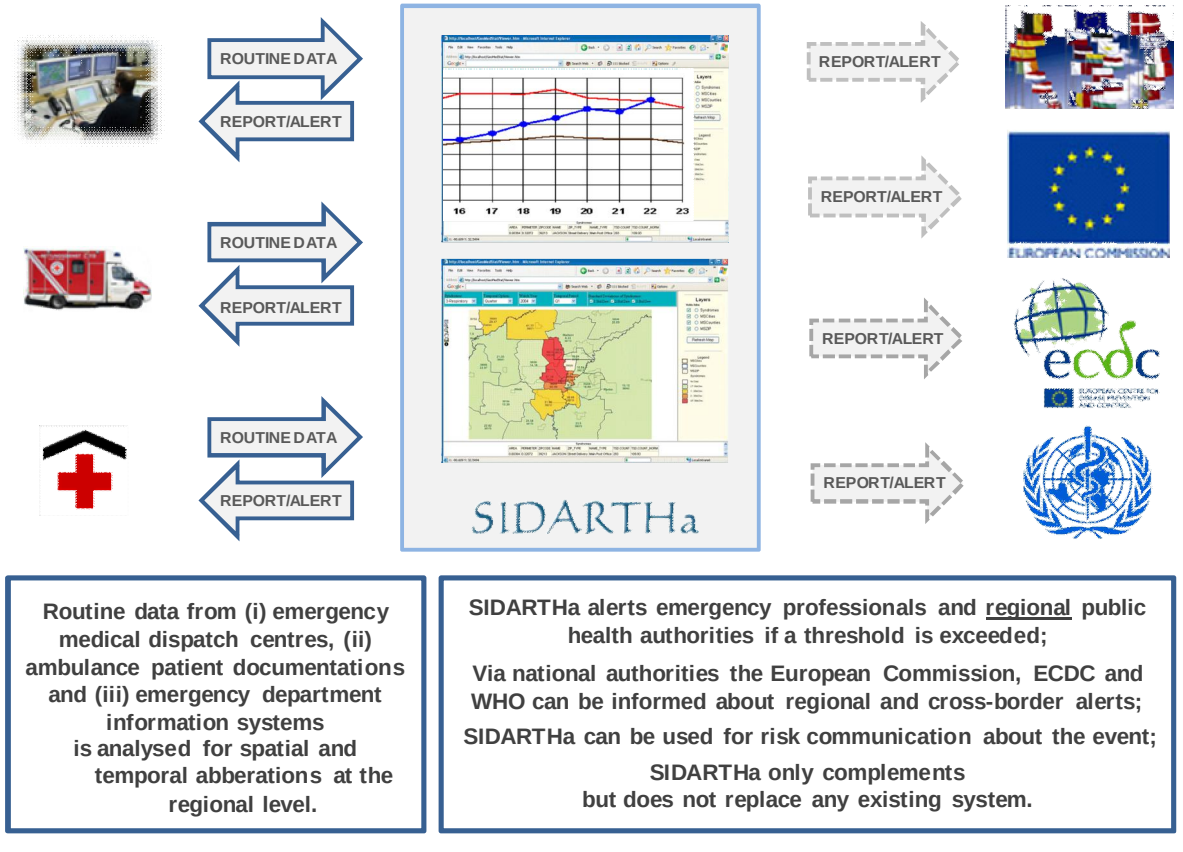

ROUTINE DATA
REPORT/ALERT
ROUTINE DATA
REPORT/ALERT
ROUTINE DATA
REPORT/ALERT
SIDARTHa
REPORT/ALERT
REPORT/ALERT
REPORT/ALERT
REPORT/ALERT
EUROPEAN COMMISSION
ecdc
Routine data from (i) emergency medical dispatch centres, (ii) ambulance patient documentations and (iii) emergency department information systems is analysed for spatial and temporal abberations at the regional level.
SIDARTHa alerts emergency professionals and regional public health authorities if a threshold is exceeded;
Via national authorities the European Commission, ECDC and WHO can be informed about regional and cross-border alerts;
SIDARTHa can be used for risk communication about the event;
SIDARTHa only complements but does not replace any existing system.

Figure 2: SIDARTHa Approach

ECDC = European Centre for Disease Prevention and Control, WHO = World Health Organization

© SIDARTHa 2010

2 Objectives & Methodology

2.1 Objectives

The main objective of this report is to provide a technical guideline or handbook for future users of the SIDARTHa syndromic surveillance system. The report describes the technical features, i.e., the system design, data flow, analysis steps, and reporting functions (Task 10 of WP 6) also including the requirements and wishes of potential future users (Task 9).

2.2 Methodology

A systematic **literature and internet review** was accomplished for existing syndromic surveillance system designs by 2 – GEOMED – DE. Next to the structured review of Medline-indexed articles that were selected as part of WP 4 (cf. Ziemann et al. 2009 (1)), a Google search was done starting with those existing systems that were described in the scientific literature identified in WP 4 and references included in these articles to other systems (groups *system descriptions* and *system evaluations*). **21 established syndromic surveillance systems** were identified. For each of these syndromic surveillance systems information was retrieved from the articles and from internet content published by the system providers covering the categories *system description*, *flow chart* (figure), *system evaluation*, *manuals*, and *miscellaneous*.

During the first **Technical Workshop** the implementation site representatives[4] and the technical unit came together with additional experts for statistical modelling and software programming to discuss the different established approaches towards system design. Based on this international state-of-the-art analysis the technical working group agreed on a system design for the SIDARTHa syndromic surveillance system. During **Technical Workshop II** the system design was revised based on the first experiences by setting up the beta-version of the system in the Spanish implementation site.

1 – UNICAN – ES works together with the company *Be Valley Technologies* which provides data analysis applications based on open source software for health institutions in Spain. This company was chosen by the Steering Committee to be the main partner for **software development** (automatic SIDARTHa syndromic surveillance application) to the SIDARTHa project group instead of having partners in each implementation site. This will allow for a standardised application fostering the acceptance of the system. Funds for this cooperation are provided as own contribution by 1 – UNICAN – ES from a local grant successfully applied for together with *Be Valley Technologies* as part of a wider cooperation project. During

[4] SIDARTHa Implementation sites: District of Kufstein, Austria; Capital Region, Denmark, County of Goeppingen, Germany, Autonomous Region Cantabria, Spain

© SIDARTHa 2010

the setup phase *Be Valley Technologies* has developed a first beta version comprising automatic data analysis with the detection algorithms, automatic alerting/reporting function (standardised report), an interactive display where further temporal and spatial analysis functions can be chosen and the option to share analysis results for risk communication and comparison.

The originally planned Delphi-type study to identify external potential future user's desires for system features was adjusted during Steering Committee Meeting V towards a more targeted approach of a short half-standardised **potential future user survey** to be distributed at special events such as European conferences. The survey covers the following aspects:

- Your Role: area of expertise/work, country, administrative level (regional, national etc.);
- Utility of Syndromic Surveillance: experience in syndromic surveillance, area of application of syndromic surveillance, top five characteristics of syndromic surveillance, reasons for non-use of syndromic surveillance;
- Design of Syndromic Surveillance Systems: Top five characteristics of the optimal syndromic surveillance system;
- Finally…: Usefulness of syndromic surveillance in Europe, age and gender of the interviewee.

The survey was distributed among 20 participants of the workshop *Syndromic Surveillance in Europe: Needed or Needless?* organised by the SIDARTHa consortium at the Second European Public Health Conference in Lodz, Poland on November 27, 2009 and at two presentations by the SIDARTHa consortium during the Third European Public Health Conference in Amsterdam, Netherlands between November 10 and 13, 2010 (23 participants). The survey can be found in Appendix 1.

The actual test and evaluation by future users during the implementation phase (WP 7) was another more practical way to adjust the system to user wishes and requirements.

© SIDARTHa 2010

3 SIDARTHa Syndromic Surveillance System Design

3.1 State-of-the-Art Syndromic Surveillance System Design

The review of system designs from the literature and internet sources resulted in a common format or template of components of a state-of-the-art syndromic surveillance system (Figure 3):

- Data Collection
- Data Classification
- Data Warehousing & Database Encapsulation
- Outbreak detection
- User Interfaces
- Communication/Network
- Implementation & Usage [Evaluation & Validation]

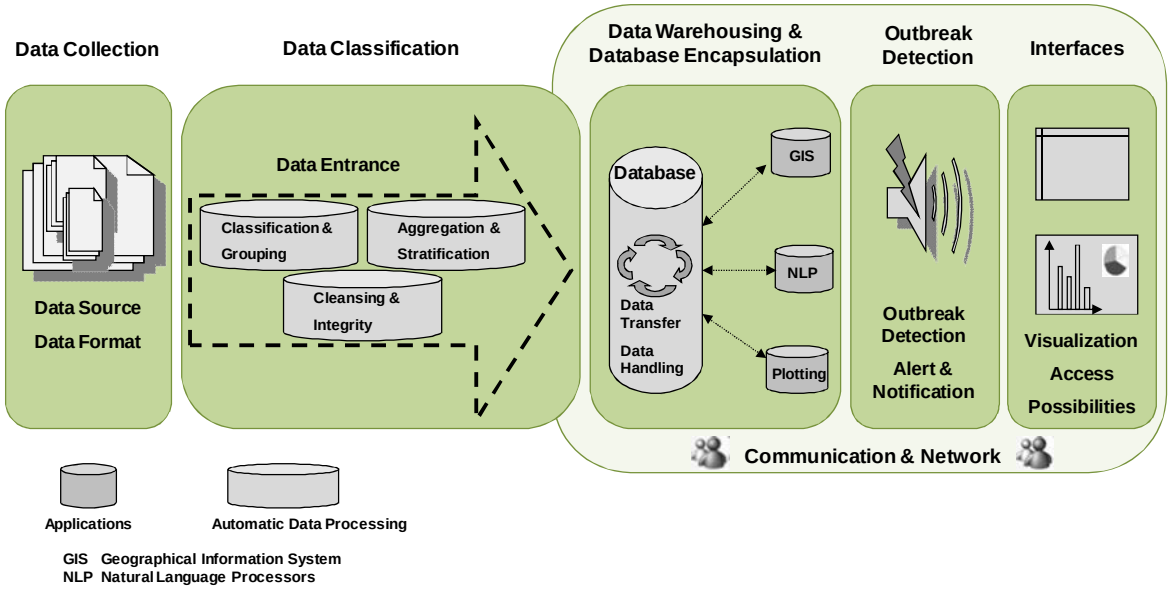


Figure 3: Components of Syndromic Surveillance Systems

© SIDARTHa 2010

3.2 SIDARTHa Syndromic Surveillance System Design

Based on the results of the state-of-the-art analysis the technical working group redesigned the first draft of the SIDARTHa syndromic system design (cf. Annex I to the Grant Agreement) to what is shown in Figure 4 and Figure 5.

The syndromic surveillance system can be installed at single emergency institutions or also at regional level where data from single emergency institutions is transferred to another or a higher institution. Baselines and thresholds must be calculated based on historical data separately for each of the emergency data sources (EMD, EMS, ED) in one region. Data on one case from different data sources should not be merged – this poses too many threats to the approach and there is no actual need (e.g., missing cases), duplicates are not affecting the approach as baselines are taking into account solely the historical demand, portability of the approach will be higher as institutions in one area do not depend on each other and raw data does not have to leave the institution/region. The emergency institution which installs SIDARTHa needs to write a translation between the own emergency data base, the individual data field names of another language and different field contents, and the SIDARTHa application. This can be done for example by creating an Extract-Transfer-Load process (ETL) which the IT administrator of the institution will be able to programme. The query will be permanently installed and selects and translates those data variables from the emergency institution's data base that the SIDARTHa application can automatically process – the standardised SIDARTHa data set (cf. SIDARTHa Coding Manual: Garcia-Castrillo Riesgo et al. 2009 (2)).

The SIDARTHa application automatically and periodically analyses the continuously updated SIDARTHa data set at a regular basis (e.g., daily) – different temporal detection algorithms will be applied providing results based on different baselines and thresholds that are calculated for the day. The application provides a simple overview if a threshold is exceeded or not (alert clock dashboard). Detailed information can be accessed and various analyses of the data can be chosen by the user. The regional administrator of the SIDARTHa system can grant access to any staff member or stakeholder to see the alerts and analysis results online.

The response to an alert will need communication and collaboration between regional public health authorities and emergency professionals. The regional public health authority can inform regional or national authorities following the established national reporting ways. Via national authorities, the European institutions and networks can be informed. The application also enables cross-border (or just cross-region) sharing of information.

A direct links to the EU Medical Intelligence System (MedISys) is explored.

© SIDARTHa 2010

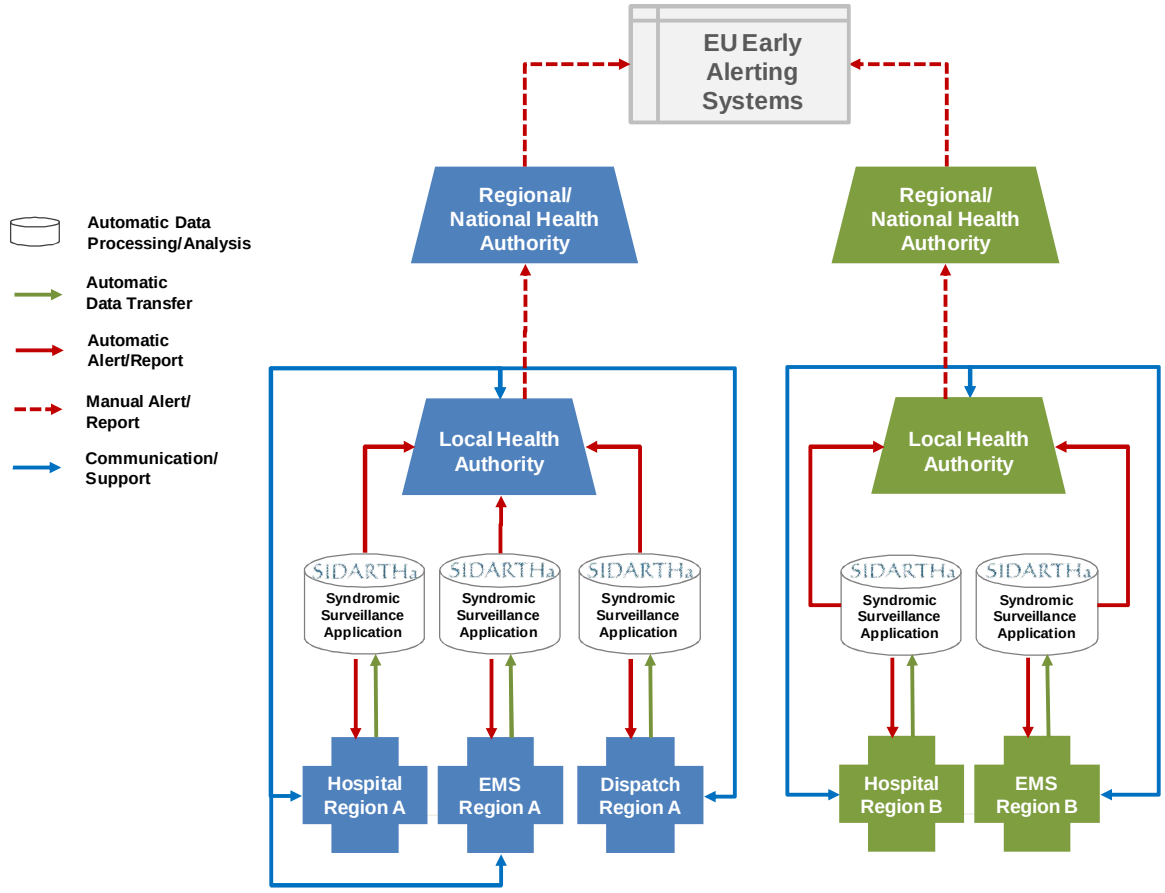


Figure 4: SIDARTHa Syndromic Surveillance System Design (Version 2)

EMS = Emergency Medical Service, EU = European Union

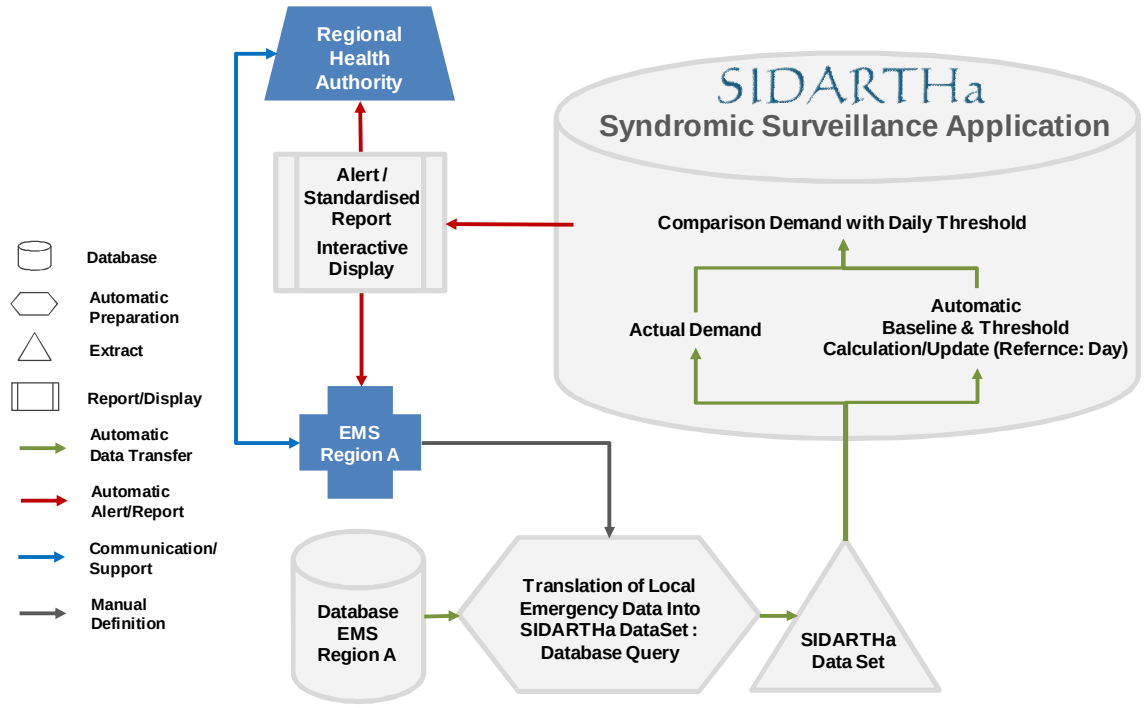


Figure 5: SIDARTHa Syndromic Surveillance System Design – Regional Level (Version 2)

EMS = Emergency Medical Service

© SIDARTHa 2010

3.3 SIDARTHa Syndromic Surveillance Application

The core of the SIDARTHa syndromic surveillance system is the automated SIDARTHa application. This piece of software enables the automated analysis of the data and display of results, production of standard reports and exchange with other users via the internet (password restricted access). The technical architecture of this application is described in the following section.

The SIDARTHa application is based on a business intelligence platform developed by *BeValley Technologies* for web-based hospital data analyses in Spain called *BeValley Health*. Appendix 2 provides an overview on the technical development of this generic fundament of the SIDARTHa application.

All information from the SIDARTHa application is stored in a data warehouse. The database of the data warehouse is designed to increase the performance in data analysis, namely that it is ready to exploit the information with available business intelligence tools.

The SIDARTHa application is based on *JBoss Seam* to ensure that no user can access data they should not. The technological architecture of the underlying *BeValley Health* platform has been based on the following components – if possible based on open source software: the web interface incorporates tools for decision support, research promotion, social networking and benchmarking. The platform has been developed following a service-oriented architecture (SOA) deployed on *JBoss AS* as application *J2EE* server. Furthermore, it has provided a portal to a dynamic interface using *JBoss RichFaces*, which allows *AJAX* and *JSF* technology to be integrated.

The automated detection algorithms were constructed to calculate the alarm limits in *R* language of *R-Project.* Algorithms included in the initial version are *ARIMA, Holt Winters Smoothing, CUSUM for normal and Poisson distributed data, StructTS, AR* and *C1, C2, C3.*

Figure 6 provides an overview on the technical architecture of the SIDARTHa application.

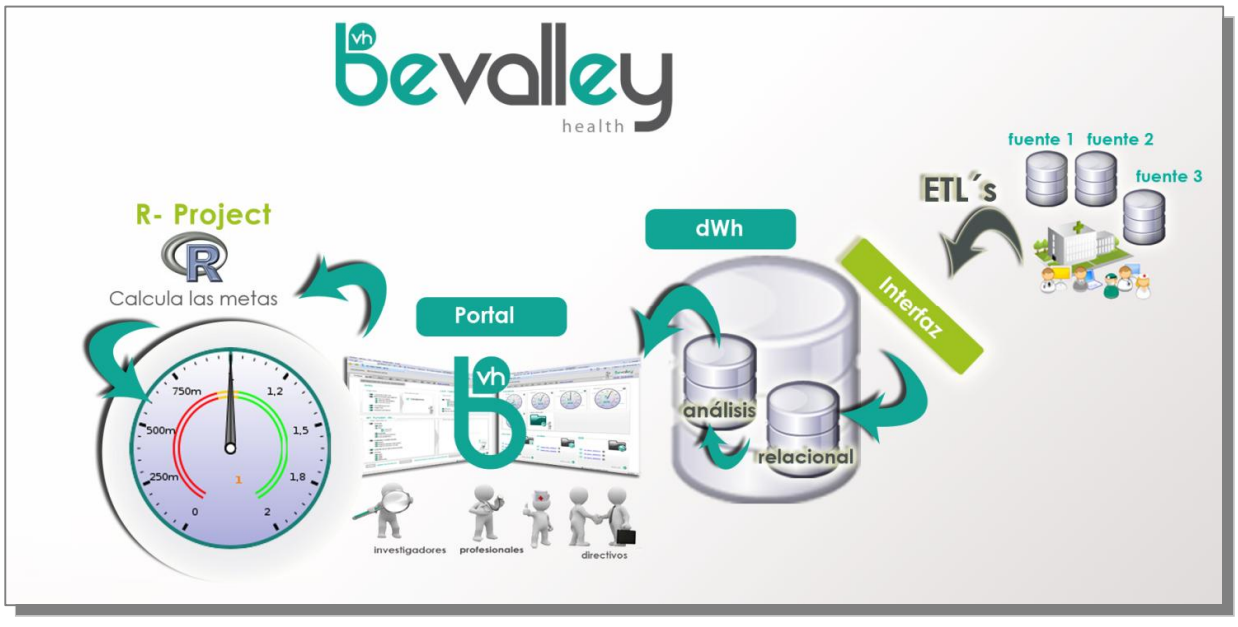


Figure 6: Technical setup of the SIDARTHa syndromic surveillance application

Bvh = BeValley Health; dWh = Data Ware House; ETL = Extract Transfer Load

© SIDARTHa 2010

3.4 Future User Requirements

The 43 interviewees were mainly working in the areas of public health practice, public health science and epidemiology/statistics. One third of the respondents have experience in syndromic surveillance using it mainly in public health practice, science and during special events/mass gatherings, not in health care practice or decision making. The experienced interviewees chose a broad range of important characteristics of syndromic surveillance, the top three being usefullnes during special events, timeliness between event and reporting and compared to traditional surveillance, and usefulness for event detection. Data quality and accuracy was seen as another crucial aspect of syndromic surveillance. Those not having experience with syndromic surveillance gave the following top four reasons: no collaboration with data providers, missing expertise, missing infrastructure, and no suitable data that is collected. The overall answer if syndromic surveillance is a useful concept was clearly answered with yes (figure 7).

The main issue raised as reason why public healt professionals do not use syndromic surveillance – the missing collaboration with data providers – is overcome in the SIDARTHa project being an initiative of the data providers themselves. This makes sure that most is made out of the available data and correct assumptions and interpretations are drawn based on the data. Nevertheless, the link between emergency care and public health authorities in the use of SIDARTHa must be strengthened in the pilot regions. SIDARTHa can also be a chance to enhance regional collaboration in public health crisis but also in every-day business.

The main uses of syndromic surveillance (special event surveillance, event detection) are covered by SIDARTHa as it applies a flexible and braod approach. Timeliness is given using a daily and automated analysis of data. Data quality is assured by standardised, automated data transfer and analysis. Being based in emerergency care institutions, SIDARTHa has the potential to be used for decision making and managing resources in these institutions as well.

The results of the survey also show the need to advertise the use of syndromic surveillance and provide knowledge and good practice experiences from experienced syndromic surveillance users to those not yet experienced. The SIDARTHa project will publish all results and provides its syndromic surveillance application free-of-charge, it will provide handbooks and recommendations for users of SIDARTHa to foster the acceptance of the system.

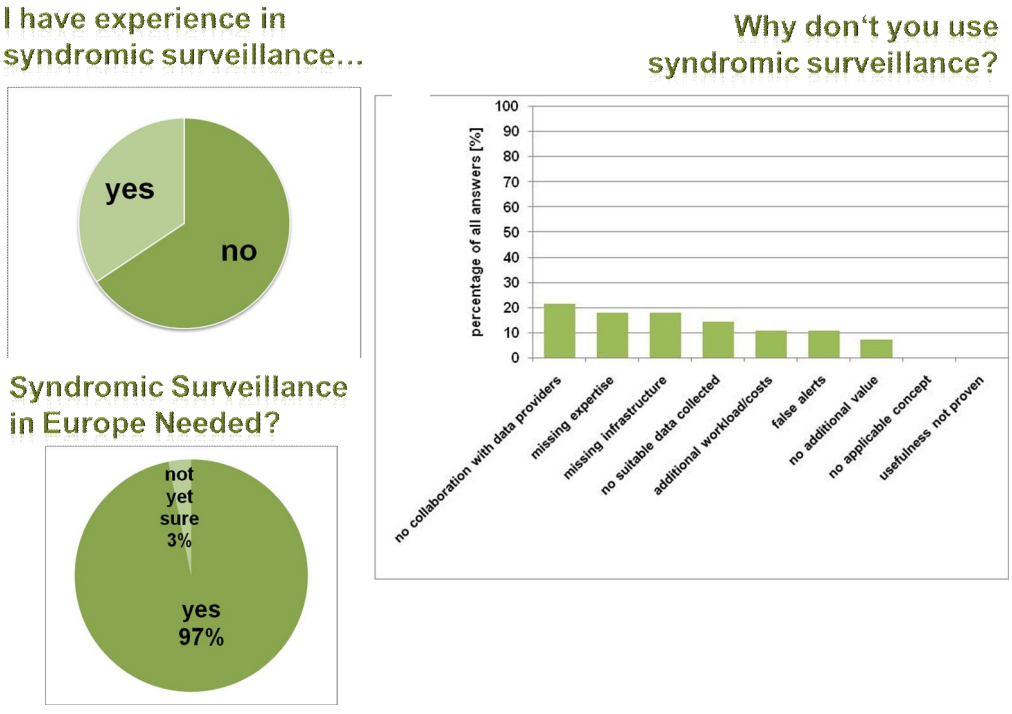


Figure 7: SIDARTHa short survey results

© SIDARTHa 2010

4 Technical Guidelines

4.1 Demo Version & Support

The SIDARTHa syndromic surveillance application is also available as **demo version** in order to be tested by potential future users before installing the software in their own institutions. The demo version uses artificial data and displays predefined analyses to show the design and functions of the software.

The demo version is accessible online and must not be downloaded. It can be accessed through the SIDARTHa website:

www.sidartha.eu/surveillancesystem/demo.htm

To use the demo version the user has to register with *BeValley Technologies* to generate a login and password to use the Demo Version for free. After having logged-in, the user must type in "demo sidartha" in the search field on top of the page. The geographical component is disabled in the demo version for confidentiality reasons.

searching machine "demo Sidartha" and you will have access to a *Demo Sidartha* data base, and two indicators: one graph with "*Respiratory Syndrome*", and the alert indicator.

On the SIDARTHa syndromic surveillance system website further **supporting documents** for users are available in the form of this user manual and the coding manual (after approval by the EU):

www.sidartha.eu/surveillancesystem.htm

Further supporting material can be found on the SIDARTHa website in the form of all work package reports or publications.

Individual support for users for questions related to the contents and the SIDARTHa approach is provided by the SIDARTHa consortium via the SIDARTHa Scientific-Technical Coordination Office which functions as central contact point.

Technical support and guidance is provided by the provider of the software, *BeValley Technologies* via video tutorials, webinars, FAQ lists, and individually via a contact link in the software.

© SIDARTHa 2010

4.2 Download & Setup

The SIDARTHa application can be downloaded under the following two links to the generic platform (BeValley) and the SIDARTHa module:

- BeValley: http://download.bevalley.com/
- SIDARTHa module: http://download.bevalley.com/sidartha/

Download and installation needs some IT expertise *BeValley Technologies* group provides support on this process to future users.

The following setup instructions by *BeValley Technologies* are targeted at IT personal in the emergency institutions:

First, you will have to download *JBoss AS*. Currently, we are using version *5.1.0.GA* although any higher version should work as well. Here is the download link:

http://sourceforge.net/projects/jboss/files/JBoss/JBoss-5.1.0.GA/jboss-5.1.0.GA-jdk6.zip/download

As the file name suggests, you will need a *JDK 6* environment. Instructions on how to install and configure the server can be found in the *JBoss Community* page.

You will also need a *MySQL 5* database server. Create a user named *bit* with password *bit* and grant it with rights for databases *bit* and *dwh*.

Populate the *bit* and *dwh* databases with the following *SQL* scripts:

- BIT -> bit.sql
- DWH -> dwh.sql

The last step is deploying the application into the running server. Download every file and copy it to the /server/default/deploy directory in the *JBoss AS* installation.

- BIT data source -> bit-mysql-ds.xml
- DWH data source -> dwh-mysql-xa-ds.xml
- BeValley health application -> health-bit-application-2010.02.1.ear
  (md5sum: ba467c317cd56eb780bfa4a966dc8a11)

There are ten predefined users: *super* who has access to all data, and another nine ranging from *user1* to *user9*, who won't be able to see anything until *super* (or another user) creates a profile for them. All passwords are equal to the login.

There aren't any health data so you will have to load it into the *dwh* database. Here is a brief explanation of the tables for the ED context (the detailed tables can be found in Appendix 3):

- *emergency*: every emergency should be reflected with exactly one record.
- *centre*: the medical centres are stored in this table.
- *postalCode*: the postal codes of the patients are stored in this table.
- *dayOfWeek*: the days of the week are stored in this table.
- *age*: the ages of the patients are stored in this table.
- *date*: the dates of the emergencies are stored in this table.
- *time*: the times of the emergencies are stored in this table.
- *department*: the departments of the hospital or EMS are stored in this table.
- *gender*: the gender of the emergencies are stored in this table.
- *icd*: the codes according to the ICD specification are stored in this table. We are currently using ICD-9-CM, as well as ICD-O-3.
- *severity*: the severities of the emergencies are stored in this table.
- *Consultation Cause*: the reasons for demanding care are stored in this table.

The integration of SIDARTHa is accomplished through an *XML* configuration file. You can update any indicator you have previously created with predictions based on previous values.

The configuration file has to be named *sidartha.xml*, and has to be placed in a directory called *etc* in the root of the same file system where you have installed *JBoss AS*.

You can see the file format in

Figure 8.

The first three parameters of each *SIDARTHa-indicator* make reference to the user that created the indicator that has to be updated, the dashboard (i.e., tab) where the indicator is located and the name of the indicator. The only requisite is that the time filter for the indicator is *From the start of day*, the prediction is estimated using aggregates for day periods

© SIDARTHa 2010

so any other time filter will lead to an erroneous prediction. However, the indicator may have any other additional filter (gender, department) that will be taken into account for calculations, thus providing a great flexibility.

Then, you can optionally add a *warning* element, a *critical* element, and as many *benchmark* elements as you want (but in this order). The value of any of these elements is the algorithm used for predicition. Valid values are: *HoltWinters*, *StructTS*, *AR*, *ARIMA, CUSUM* and *C1, C2, C3*.

For the prediction, we use the values of the last 250 weeks (1750 days), not including today. This parameter will be configurable in future releases. VERY IMPORTANT: please make sure that you populate the *date* table with values for all these days, even if there are no corresponding facts. This is the only way to tell the analysis tool that the aggregate for that day has been zero.

The update process is intended to be executed periodically according to a configurable schedule but for this first implementation you will have to update the indicators manually. Please use the following URL:

http://server:8080/health-bit-sidartha

in which *server* is the host name of the server where *bVh* is installed.

```
<?xml version="1.0" encoding="UTF-8"?>

<sidartha xmlns="http://health.bevalley.com/sidartha"
        xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
        xsi:schemaLocation="http://health.bevalley.com/sidartha
sidartha.xsd">
        <sidartha-indicator>
                <user>user2</user>
                <dashboard>Panel 1</dashboard>
                <indicator>sidartha test</indicator>
                <warning>ar</warning>
                <critical>StructTS</critical>
                <benchmark>arima</benchmark>
        </sidartha-indicator>
        <sidartha-indicator>
                <user>user1</user>
                <dashboard>Panel 2</dashboard>
                <indicator>sidartha test</indicator>
                <critical>arima</critical>
                <benchmark>c123</benchmark>
                <benchmark>HoltWinters</benchmark>
        </sidartha-indicator>
</sidartha>
```

Figure 8: SIDARTHa application: file format

© SIDARTHa 2010

4.2 User Manual

4.2.1 SIDARTHa Local Health Surveillance System (LHSS)

The Wall

After login in the tool provides different options from the main page or the Wall:

- User profile including an overview on contacts and indicator groups (right)
- Upload of new data (top)
- Search machine (top centre)
- All data bases and analyses (left)
- Message board (top left)
- Help and FAQ (bottom)

The user profile shows contacts, databases, and indicator groups (labels) covering e.g. all analyses related to a certain data source.

The left central area contains all data bases, indicators, tables or graphs generated by the user or shared by other users. Each item is displayed in a box containing also further information or comments.

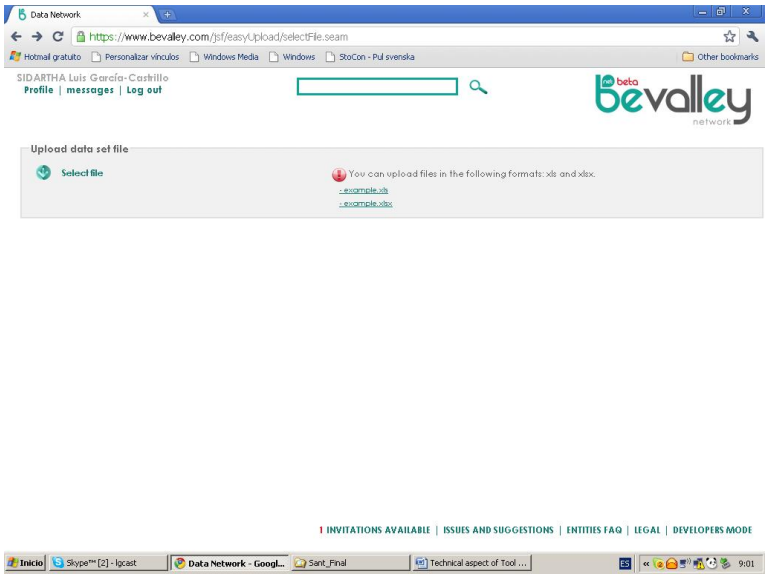
All functions are accessible from the Wall:

- Data upload
- Generation or change of indicators/analyses
- Change of alert thresholds
- Generation of labels

Data Upload

Data files can be uploaded manually or automatically using an ETL application. In the current version of the application, the uploaded files must be Excel files. They can contain an unlimited number of cases. The data files should follow the SIDARTHa Standard Data Set (cf. Coding Manual; Garcia-Castrillo Riesgo et al. 2009 (2)) and contain the minimum data fields of date case, and geographical information for map representation. Each row represents a case.

© SIDARTHa 2010


Data Network
https://www.bevalley.com/jsf/easyUpload/selectFile.seam
SIDARTHA Luis García-Castrillo
Profile | messages | Log out
beta bevalley network
Upload data set file
Select file
You can upload files in the following formats: xls and xlsx.
- example.xls
- example.xlsx
1 INVITATIONS AVAILABLE | ISSUES AND SUGGESTIONS | ENTITIES FAQ | LEGAL | DEVELOPERS MODE

It is recommended to check the uploaded data set using the function *Management of data set*. Successfully uploaded data sets are accessible from the Wall.

Using Data

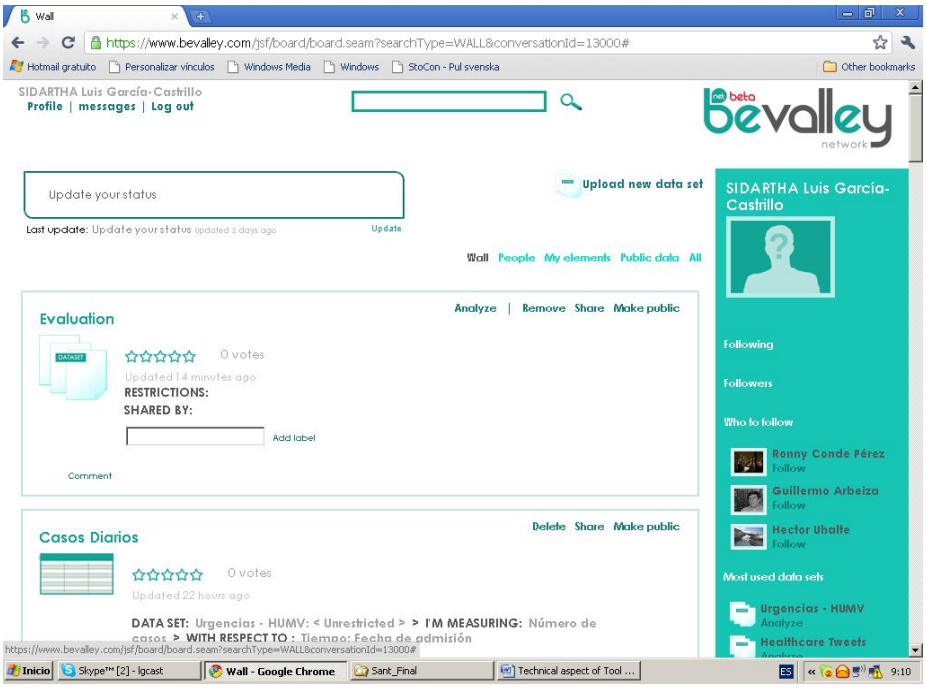
For an uploaded data base several actions can be chose:

- *Analyze*
- *Share*
- *Delete*
- *Make Public*

Or one can click on the database to change its name or access the option "*Management of data set*."

Another option is to let the database become part of a group using "*Add a label*". Labels facilitate navigation by grouping diferent elements (data bases, indicators, graphs, or maps) under one label.

The labels can be selected from the "*Wall*" in the green box on the right.



© SIDARTHa 2010

Analysis / Indicator generation

This function allows you to create indicators for specific syndromes, set alerts, and generate graphs, tables or maps.

Clicking on "*Analyze*" in the box of a data base a selection of analyses are presented: Graphs, tables, indicators, indicator with predictive analysis and map representation.

After having chosen the analysis option the characteristics are defined in the next step: static filter (e.g., specific syndrome), measure (e.g., number of cases), time filter (e.g., today), benchmark (detection algorithm and alert threshold).

Wall
https://www.bevalley.com/jsf/board/board.seam?searchType=WALL&conversationId=13000#
SIDARTHA Luis García-Castrillo
Profile | messages | Log out
Cancel
Graph
Represent data with line or bar graphs, or pie charts.
Time graph
Represent the evolution in time of one or more variables and compare them with other periods. The date which the analysis is regarding to (reference date) can be modified.
Table
Represent data in tables which makes it possible to go deeper until reach the original data. Besides, data can be exported to different formats.
Indicator
Represent the current value for a measure (reference date can be modified). You can set comparison values based on other periods or fixed benchmarks.
Prediction indicator
Can generate alarms to represent the value of the previous day compared with predicted values for that day and the application of a statistical margin of error
Update your status
Evaluation
Casos Diarios
RESTRICTIONS:
SHARED BY:
DATA SET: Urgencias - HUMV: < Unrestricted > I'M MEASURING: Número de casos > WITH RESPECT TO : Tiempo: Fecha de admisión
Following
Followers
Who to follow
Most used data sets
Urgencias - HUMV
Healthcare Tweets

Indicator Generation

Two types of indicators are offered; general indicators with basic analytical information (means, sum etc) and predictive indicators. The latter are useful to compare predicted with actual values, and set thresholds for alerts.

At the *Indicator* screen select *Static Filter* and chose the variables that generate the desired syndrome. Select an individual variable or a combination of multiple variables that defines the desired syndrome in the left green window. *Accept* the selection.

Select *Measure* for different values, e.g., number of cases (rows). *Accept* the selection.

For the *Time filter* the previous day is set by default but specific dates and time periods can be chosen. *Accept* the selection.

Select *Benchmark* to change the detection algorithm and change the alert threshold. *Accept* the selection.

Any combination of detection algorithms can be chosen. One can set *limits* for one algorithm and chose another as *benchmark* (e.g., *CUSUM* as upper limit and *C1, C2, C3* as benchmark).

It always has to be defined if high or low values generate the alert.

A wide range of thresholds can be selected:

- Fixed values
- Benchmark with historical data
- Prediction benchmark with a selection of algorithms

© SIDARTHa 2010

If a data set with a high number of cases over a longer period is analysed (i.e., total number of cases) it is recommended to use algorithms that take into consideration seasonal variation (if enough historical data is available) such as *ARIMA, Struct Ts, or Holt Winters*.

For smaller or more specific data sets, e.g., syndromes, it is recommended to use *CUSUM* or *C1, C2, C3*.

Save the newly defined indicator by clicking "*Save and exit*".

Modifications can be done to any of the components later on. Choose *Discard,* if you do not want to save the indicator.

Once the indicator has been saved it is updated automatically. In the box below the figure the last updates are stated.

© SIDARTHa 2010

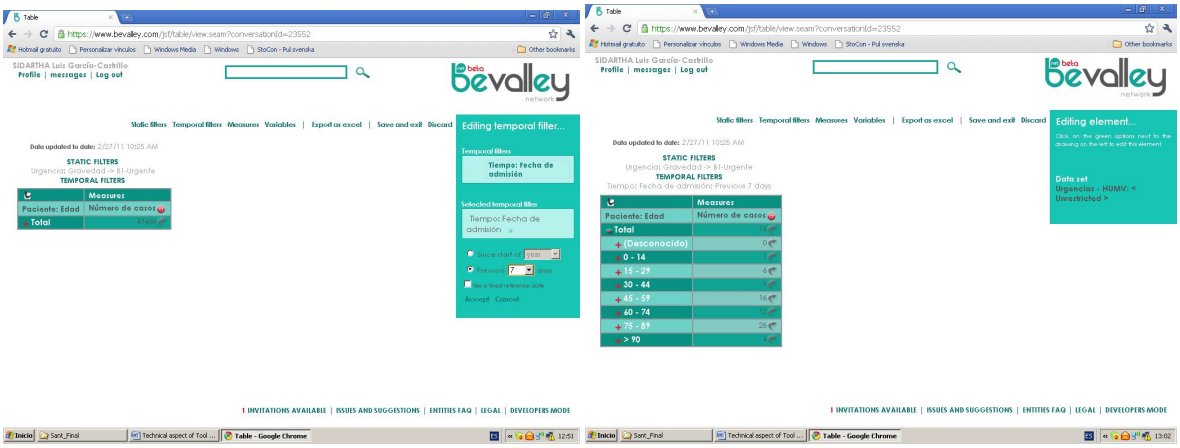
Tables

Tables can be used to do further analyse a specific event, i.e., a day with an alert. It is possible to export data as excel or pdf files.

From the *Wall* select the database to be used, click *Analyze* and then *Tables,* a table without any filter or specification is generated. As for graphs, maps and indicators, static filters, temporal filters, and measures can be defined.

Over the table the defined characteristics are displayed.

Before you leave the screen save the table by clicking *Save and exit*. Modifications can be done to any of the components later. Click "*Discard"* if the table should not be saved.



Maps

In this version of the SIDARTHa application, *Google Maps* is used to represent the spatial distribution of cases in absolute numbers (if spatial information is included in the uploaded data base).

© SIDARTHa 2010

4.2.2 SIDARTHa European Health Surveillance System (EHSS)

The SIDARTHa application is designed to facilitate exchange between different users of the system. The extent to which results or data are exchanged is decided by the owner of the information.

Information, databases, graphs, tables or indicators can be shared to any other registered user of the SIDARTHa application.

There are two levels of sharing information: the public or specific users. To share information *Share* can be selected in the respective box of one indicator or data base.

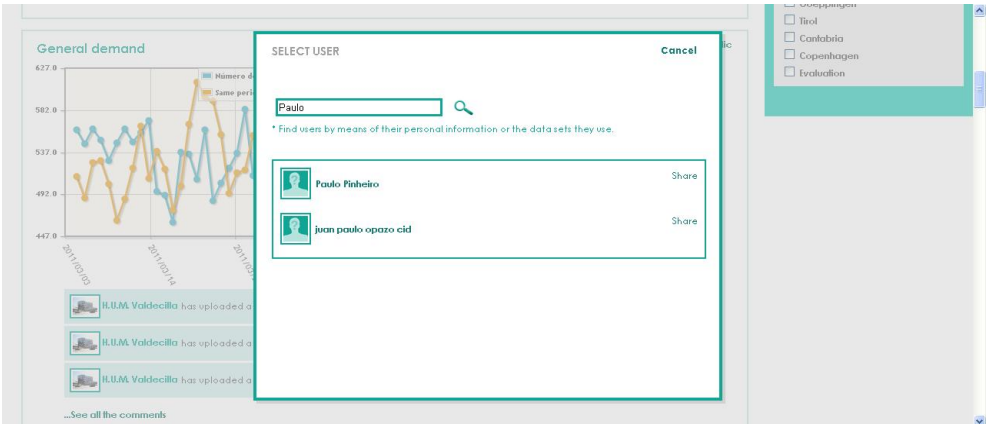
A new window with a search tool allows you to choose the registered user to share the information with.

When clicking on an indicator the list of persons this indicator has been shared with appears on the right (green information box). Here, sharing of this indicator can be disabled for a selected user.

When sharing an indicator, table or map, the user can only see it. When a data base is shared the other user can analyse the data and generate new indicators.

When choosing *Make public* the indicator or data base is shared with any member of the BeValley Network (beyond the SIDARTHa network).

Users are informed when information was shared with them via the message board. Access to messages is possible via the *Wall*.



© SIDARTHa 2010

References

1. Ziemann A, Krafft T, Garcia-Castrillo Riesgo L, Fischer M, Krämer A, Lippert F, Vergeiner G for the SIDARTHa project group (eds.). Early Detection of Health Threats: International State-of-the-Art & European Context. Results from the SIDARTHa project. 2009, Bad Honnef.

2. Garcia-Castrillo Riesgo L, Krafft T, Ziemann A, Fischer M, Lippert F, Vergeiner G, Krämer A, Brand H for the SIDARTHa project group (eds.). The SIDARTHa Coding Manual – How to generate syndromes based on routinely collected emergency care data for the European syndromic surveillance system SIDARTHa. 2009, Bad Honnef.

© SIDARTHa 2010

Appendix

© SIDARTHa 2010

Appendix 1 SIDARTHa Short Survey on User Requirements

© SIDARTHa 2010

SIDARTHa
European Emergency Data-based Syndromic Surveillance System

This project is co-funded by the European Commission,
DG Health & Consumers (2007208).

GEOMED RESEARCH UC UNIVERSIDAD DE CANTABRIA

www.sidartha.eu

Survey
Syndromic Surveillance in Europe: Needed or Needless?

Background: This survey is set within the framework of the European Commission co-funded project SIDARTHa – European Emergency Data-based Syndromic Surveillance System (Grant Agreement No. 2007208; project time: June 2008 – December 2010). The aim of the project is to contribute to and to enhance the generic preparedness of the health sector for communicable as well as non-communicable health threats and emergencies by conceptualising, developing, implementing/testing and evaluating an innovative syndromic surveillance system using indicators derived from routinely collected emergency medical services (prehospital) and emergency department (hospital) data. The SIDARTHa syndromic surveillance system targets at the local/regional level and does not replace any existing surveillance structures.

Survey Purpose: Assessment of requirements and needs for syndromic surveillance in Europe. Your replies in this survey will directly be incorporated in the design of the SIDARTHa syndromic surveillance system. The SIDARTHa syndromic surveillance software application will be available for implementation across Europe free-of-charge after the project time.

Definition: Syndromic surveillance is understood in this survey as
"an investigational approach where health department staff, assisted by automated data acquisition and generation of statistical signals, monitor disease indicators continually (real-time) or at least daily (near real-time) to detect outbreaks of diseases earlier and more completely than might otherwise be possible with traditional public health methods (e.g., by reportable disease surveillance and telephone consultation)"

We very much appreciate your support in completing and returning this survey to the SIDARTHa team!

Thank you very much!

YOUR ROLE

1 I am mainly involved in…

1.1.1 ☐ public health
1.1.2 ☐ health care
1.1.3 ☐ epidemiology/statistics
1.1.4 ☐ other

1.2.1 ☐ scientifically
1.2.2 ☐ professionally
1.2.3 ☐ politically

2 I am mainly working in the country of

_____.

3 I am mainly/currently working at the following level:

3.1 ☐ supranational level
3.2 ☐ national level
3.3 ☐ subnational/state/regional level
3.4 ☐ local level

© SIDARTHa 2010

UTILITY OF SYNDROMIC SURVEILLANCE

4 I have experience in syndromic surveillance:

☐ no → → → → please go to question 6, page 3

☐ yes

IF YES:

I use syndromic surveillance… (please tick all that apply)

4.1 ☐ in every-day public health practice.

4.2 ☐ in every-day health care practice.

4.3 ☐ in every-day decision making.

4.4 ☐ in scientific/methodological work.

4.5 ☐ in special events/mass gatherings/disasters (situational awareness).

4.6 ☐ other:_____

5 To me, the following 5 characteristics of syndromic surveillance are most important (please tick up to five boxes)

5.1 ☐ Timeliness between event and reporting

5.2 ☐ Timeliness compared to laboratory-based or sentinel surveillance

5.3 ☐ Data quality/accuracy

5.4 ☐ Validity (Predictive Value)

5.5 ☐ Flexibility of the approach (rapid setup, adjustment of data analysis, data source selection, utility for different public health questions)

5.6 ☐ Cost effectiveness

5.7 ☐ Low workload

5.8 ☐ Usefulness for every-day surveillance work as supplement to traditional surveillance

5.9 ☐ Usefulness for surveillance of seasonal outbreaks

5.10 ☐ Usefulness for surveillance for detecting (known or unknown) public health threats

5.11 ☐ Usefulness for surveillance during special events (mass gatherings, natural disasters)

5.12 ☐ Close collaboration between public health and health care provision

5.13 ☐ other: _____

© SIDARTHa 2010

IF NO:

6 I do not use syndromic surveillance because…
(please tick up to five boxes)

6.1 ☐ no applicable concept is available in the European context.

6.2 ☐ the usefulness of syndromic surveillance is not proven (for the European context).

6.3 ☐ no additional value is expected compared to existing surveillance structures.

6.4 ☐ no data suitable for syndromic surveillance are collected.

6.5 ☐ no collaboration exists with providers of suitable data.

6.6 ☐ expertise/knowledge for setup/maintenance is missing.

6.7 ☐ technical infrastructure is missing.

6.8 ☐ additional workload/costs are produced.

6.9 ☐ to many false alerts are produced.

6.10 ☐ other:_____

DESIGN OF SYNDROMIC SURVEILLANCE SYSTEMS

7 If you were developing the optimal syndromic surveillance system…
(please tick up to five boxes)

7.1 ☐ I would only be informed about the detection of an event.

7.2 ☐ I would be informed regularly independent of detection of an event, e.g., daily.

7.3 ☐ I would automatically receive standardised short reports by email/fax etc. on an event.

7.4 ☐ I would want to actively analyse the syndromic surveillance data including historical events.

7.5 ☐ I would include a virtual platform for communication of all stakeholders during an event.

7.6 ☐ I would include a cross-region or cross-border analysis of events.

7.7 ☐ I would enhance collaboration between public health, health care provision and non-health stakeholders.

7.8 ☐ other:_____

FINALLY…

8 Syndromic Surveillance in Europe is…

☐ needed ☐ not needed

9 I am…

☐ male ☐ female

10 My age is…

10.1 ☐ <25 10.2 ☐ 25-40
10.3 ☐ 41-65 10.4 ☐ > 65

© SIDARTHa 2010

Appendix 2 Technical Development of BeValley Health

Project background

We want you to know where this initiative comes from and how it has been developing to the current situation, in which we will show it to you:

Starting point: EXPERIMENTAL DEVELOPMENT IN THE AREA OF SURGICAL INTERVENSIONS

As a starting point, it conducts an experimental development project involving the construction of a prototype of a technological tool (prototype "QUIVER") created for the analysis and exploitation of the surgical data, which was funded by be Valley Technologies in collaboration with the Society Cantabria Regional Development (Sodercan) through Invesnova plan. The project was implemented in the main hospital in Cantabria, Hospital Universitario Marqués de Valdecilla (Humvees), and also with the cooperation of the University of Cantabria (UC).

FEASIBILITY STUDY (1): preparatory for two simultaneous experimental development projects

Following the completion of this first tool (prototype ¨QUIVER¨) and again in collaboration with the HUMVs and its Institute for Training and Research (IFIMAV), be Valley Health Technologies supports and conducts a feasibility of preparatory character for the implementation of two experimental development projects which to be performed in a single social space on innovation (living lab).

EXPERIMENTAL DEVELOPMENT IN THE LAB OF LIVING of HUMV; IFIMAV AND 35 CAP

(PUINV Prototype and CMIBI Prototype)

These experimental development projects were the basis for the realization of two prototypes of tools that integrate and BI technologies based on Open Source to collect and exploit the living lab data (data from the HUMV, IFIMAV and 35 primary care centers in Cantabria). The first prototypes (PUINV prototype) sought to increase and improve research capacity, whereas the prototype CMIBI was aimed at supporting decision making for health professionals.

In both cases, the funding was provided by the company Be Valley Technologies, the HUMV and Sodercan through a Technology Governance Plan regionally.

The experience in the use of a social space of the innovation, together with the experience of HUMV and IFIMAV, allowed the developed prototypes to be situated at the forefront of the state of the art technology, allowing solving big limitations in which health professionals meet in their process of adaptation to the information society:

- Difficulties with management-oriented operation to determine the quality and usefulness of huge amounts and dispersed data

© SIDARTHa 2010

- Proliferation of many BI (Business Intelligence) technological tools to be selected and combined to obtain a reliable operation of the data (systems Balanced Scorecard; GIS geographic representation, reporting systems, OLAP systems, among others.)

- Limited access to information associated with the high costs of available technology (marketed by user licenses.

- Lack of standardization at all levels (data layer, business layer, etc.).

FEASIBILITY STUDY (2). Preparatory to experimental development for the construction of bVh 2.0

At this point, we find that there are number of needs that motivate a feasibility study preparatory for the experimental development of a prototype (prototype BVH 2.0) that enables decisions making in a framework where users enter their own data, define their trees elements (indicators, etc.), exchange their knowledge of business and may access to the function of the permissions according to the exploitation of all available data in the system. We considerate that, this is an extraordinary project which in its operation ensures the accessibility of all professionals (doctors, researchers, managers, etc.) to information and enables an environment in which users themselves evolve the model based on specific needs. Moreover, the prototype has been developed in the environment of a living lab in which users would have shown that the factors influencing the same.

In particular, the needs that have led us to evolve the model were:

- Lack of standardization of criteria for defining the indicators of an organization or professional, which on the other hand complicates information sharing and public health benchmarking is difficult to carry out. Moreover, it offers the opportunity to propose a unique space for discussion, the exchange of information and data for all public health professionals.

- So far the professionals who are involved in benchmarking networks, improve their decision-making processes by providing aggregated data that may have been estimated with different criteria, and therefore would not be reliable in the comparative process. In this context, it becomes apparent the usefulness of collecting the raw data and then homogenize the technological tool, to offer each professional the information in demand for each respondent according to the indicators that they themselves define in an authorized interface for this purpose.

- There are many working groups, both management and research that develop technology projects that normally they are not used for lack of data or because there are no mediators to ensure their maintenance and evolution. In this scenario, bVh 2.0 enables ways so that users can include the developed components in the tool. Thus, they will have support, development and data, meanwhile contributing to the same technological upgrading.

Characteristics

bVh 2.0 is a System integration, analysis and data mining and the dissemination and knowledge transfer (benchmarking among practitioners), which supports the decision making and research in the health sector.

Knowledge of business:

The Network has been developed by and for healthcare professionals. Initially, the collaboration with the layers of directors, professional and research institutions that currently belong to the Network, aimed at acquiring a better understanding of business, and allowed us to cut the real needs of all users in order to provide an appropriate solution to them. Although, later, we noticed that one of the

© SIDARTHa 2010

fundamental values of the model is that each entity and the user can add their own knowledge of business. To do this, the trees of definition of element definition were generated (indicators, components, data, etc.) so that all knowledge that is contributed by users becomes well organized in the lists through the implementation of algorithms to manage the same function in terms of use. Thus, de facto standards are created in addition to the officials.

Data processing:

BVH counts on a user interface to ensure the automatic collection of data for each entity, which was provided by ETL processes (extraction, transformation and loading of data). This interface has been built based on knowledge of the sources that give us the experience of living lab. Though, it is open to every entity, the users could suggest new fields for data which are not contemplated and that similarly, might be interesting . Furthermore, the data collected always raw data, so that they can be exploited to the indicators defined by each user. Once they are incorporated into the data to BVH, they are stored in the Data Warehouse (DWH), which in turn will be divided into two; operational DWH, fueled by the raw data, and DWH of analysis, on which it perform the exploitation of data with tools of the system. Note that the existence of an operational database enables the exploitation of raw data. Equally important is the fact that they examine the application of artificial intelligence (neural networks) to improve the quality of the data.

Codifying and construction of the technological tool:

The technological solution implemented business intelligence, has been constructed by integrating the following Open Source tools to work as a unique software: Database or Data Warehouse; artificial intelligence and data mining (WEKA), multidimensional data server (Mondrian); Reporting (Pentaho Reports) Balanced Scorecard (Pentaho Dashboards) and Information System geographic representation or GIS (Google Maps), system management operational database (phpMyAdmin) and OS (Ubuntu Linux).

Implementation:

The criteria of value when implementing the model has been the guarantee of accessibility to all users (managers, researchers and practitioners) without restriction of space time. To do this, it has resorted to the use of accessible technologies via web (WOA). While each entity that is wished can have the software installed on their own servers.

State of the art

2.0 BVH largely exceeds the current state of the art

State of the art (I)

- Lack of technological tools which are capable of exploiting the big amount of existing health data, either by the lack of technological capabilities or lack of functionality required for their orientation to the field of health management research or health sciences.

- Lack of integrated technology tools to identify and correct for semi-automatically by data mining systems, the quality of different sources.

- Lack of Open Source technology tools for integration with business knowledge of health (coding, data dictionaries, etc.) included.

- Proliferation of many technological tools BI (Business Intelligence) Open Source disjointed to be selected and combined to obtain a reliable operation of the data (systems Balanced Scorecard; GIS geographic representation; Reporting systems, OLAP systems; among others.).

© SIDARTHa 2010

- Lack of technological tools that give users unlimited access to BI exploitation. All existing BI solutions are architecturally designed to provide support to decision making to a low number of users (layer directive).

- Absent in the open source world of certain essential components for the exploitation of information in healthcare, both as a management-oriented research, so there is a associated limit of the high costs of available technology (marketed by licensing user).

- Lack of standardization at all levels (data layer, business layer, etc.).

These points were exceeded with the first prototype to be Valley health (BVH Prototype). Subsequently, it was found that a number of needs that motivate the development of an improved system (BVH 2.0) for taking decisions within a framework where the users enter their data, have a central repository of information, exchange knowledge of business, and could access, depending on permissions, the exploitation of all available data. Regarding the state of art, noting that is still applicable today and that is to rise to 2.0 BVH as far as possible.

State of the art (II)

- Lack of technological systems, thanks to architectural design, are capable of addressing issues of confidentiality of health authorities.

- Lack of technological systems for the homogenization of criteria defining health indicators (at individual or organizational).

- Lack of technological systems-oriented health information exchange and business knowledge and management of health at all users of an organization.

- Lack of technological systems with potential to support benchmarking health networks of global level

- Lack of technological systems that allow the performance of benchmarking health from the raw data from operational sources of health institutions (Today Benchmarking networks work without technological support, and from the aggregated indicators, previously calculated, they are not reliable in the comparison process).

- Lack of technological systems that can support the construction of technological components for third parties, with guaranteed support and accommodation, using a central repository of health information and access to specific functions of business intelligence (including data mining and artificial intelligence.

- Lack of flexible technological systems that can generate public health alarms from the routine operational sources and health institutions (alarms syndromic).

- Absence of a segmented platform, scientific and objective where healthcare providers, especially pharmacists, can advertise their prescription drugs and replace their commercial (health visitors).

- Absence of a platform that allows predictive models and hidden patterns searches from data stored in one central repository of health information systems using data mining and Artificial Intelligence.

- Absence of a technological tool that claims to be a viewer of single electronic medical record globally.

- There is a need for a global platform with clinical data and tools focused on the investigation.

© SIDARTHa 2010

Advantages

bVh 2.0 allows to have a real Health Observatory

Key features

- BVH 2.0 integrates and homogenizes the data using the available technology to facilitate their exploitation.
- It involves the introduction of an exploitation culture and increase the profitability of the information in the health sector, which contributes to the shift to organizational excellence, and that each person can have the information anytime they need. Thus it will encourage access to better health. It also promotes the sharing of ways to exploit the information for management and research.
- Advances to the democratization of information, ensuring access of any user of an organization (layer directive, professional layer, layer research) to any data (e.g. any professional can now have their own scorecard)
- to ensure the support of decision making based on objective data (allowing the reduction of waiting lists, migration assistance, etc.) and encouraging the making of comparisons in a traditionally opaque sector (benchmarking among professionals).
- to introduce the future internet in the health sector by enabling a space that evolves, thanks to the users who ensure their suitability to the specific needs of the sector.
- to allow to give support to a multitude technology projects, that in other words would not evolve because of the lack of resources for maintenance and data availability.

Economic benefits

POINTER SYSTEM OF R & D

Is a great innovation because it is based on the construction of a single technological solution Open Source (without licenses or units) available for many entities and users, and also develops and deploys technologies, applications, services and content of the Society of Information and the future internet.

POSITIONING ON THE INTERNET OF THE FUTURE

bVh integrates the internet for people, services and content; encouraging the development of the information society and knowledge and allowing us to be equipped with most developed countries in this field. In the system, the users themselves who generate content and make the model evolved according to their request. Furthermore, it evolves and improves collaborative work, creating a virtual community.
On the other hand, is integrated in the field of Internet services with a platform WOA (Web Oriented Architecture).

RESPONSIBLE FOR THE DEVELOPMENT TECHNOLOGY SECTOR

The projects aims to change the way we use the technology to exploit the information in the health sector, from a traditional model in which each entity implants and uses various technological tools to a framework that provides a unique technological tool for infinite users. We further believe that technological developments ensures the health sector because, first, it brings the future internet closer to this field and on the other hand, being based on Open Source technologies, the progress will be in parallel with the evolution in order to be happened for each of these applications that are useful to the technological tool that will support the project.

© SIDARTHa 2010

TO DEVELOP AND TO IMPROVE COMPETITIVENESS AND PROFITABILITY IN THE FIELD

It favors the increase of competitiveness (introducing a new model where there is a progress in which the administrator becomes a client) and productivity in the health sector (allowing research on management; approaching the scoreboard from one organization to all levels of users; encouraging benchmarking processes, etc.). Besides, it improves the innovation and creativity in organizations promoting the development of managerial skills for decision making.

Promotes scientific research - HEALTH

It promotes scientific research in the health sector and especially translational research which in case of sharing the data and knowledge of business, not only promotes research but also its application to a clinical practice (updating and exploitation of investigation results).

To drive democratization in access to information

It promotes democratization of data, favoring the equality in access to information, as well as intensifying and generalizing the use of ICT in the primary sector (health) or secondary (consulting, etc) in which the project ¨actúa de modo tractor¨

Roadmap

Date: 01/05/2009
Version: 2009.05 [Codename: Espinete] COMPLETE

Description:

- User Interface oriented technology division.
- Blocks analysis of hospitalization data validated by the users of the living-lab:
  - Hospitalization (episodes, sub-episodes of service, waiting lists, visit sessions)
  - Surgical interventions (interventions, waiting lists, surgical sessions)
  - Diagnostic techniques (techniques, waiting lists, technical sessions)
  - Emergencies
  - Primary care episodes
  - Primary care pharmaceutical prescription
  - Primary care immunizations

Technology modules
- OLAP analysis (First stable version)
- Control panel system and indicator generation from OLAP (First stable version)
- GIS Analysis (first stable version based on MapXtreme)

© SIDARTHa 2010

Date: 01/06/2009
Version: 2009.06 [codename: Caponata] COMPLETE

Description:

- User Interface oriented to end users.

> The same blocks *of data analysis with the primary care data through living-lab users.*
> *Technology modules:*

• Ability to export records to SPSS from searches made with the OLAP system.
• Incorporation of PHPMyAdmin to analyze the Operational Data Warehouse

Date: 03/07/2009
Version: 2009.07 [codename: Nabucodonosorcito] COMPLETE

Description:
- Ability to generate reports from the leaves and the views of users
- A support for OLAP query generation
- A support for Geographic Information Systems (GIS) generation

Date: 21/09/2009
Version: 2009.09 [codename: Juan Olvido] COMPLETE
Description:

> A management of Independent dimension "time"
> Incorporation of obligatory filters
> Incorporation of targets and references to previous periods

Date: 01/01/2010
Version: 2010.01 [codename: Perezgil]

Description:

> A management of Independent dimension "entity"
> Incorporation of functionalities for the relationship between entities
> Implementation of the benchmarking system

Date: 01/02/2010 Version: 2010.02 [codename: Conde Draco]

Description:

Transformation of analysis network into a complete SOCIAL NETWORK analysis and comparison

Date: 01/03/2010
Versio: 2010.03 [codename: Sheriff Coco]

Descripción:
The completion of the API for the integration of the third party applications.

© SIDARTHa 2010

Appendix 3 SIDARTHa Application Setup: Data base contents

- *emergency.* every emergency should be reflected with exactly one record in this table. It has the following columns:

column	description
Id	ID of the emergency. It is marked as auto increment, so you should leave it blank when inserting the record and let the database assing the ID for you.
stayTime	Period that the patient stays in the hospital.
centreId	ID of the centre where the emergency happens. It references the *centre* table.
postalCodeId	ID of the centre where the emergency happens. It references the *postalCode* table.
admissionDayOfWeekId	ID of the day of week for the emergency. It references the *dayOfWeek* table.
ageId	ID of the age of the patient. It references the *age* table.
admissionDateId	ID of the admission date for the emergency. It references the *date* table.
admissionTimeId	ID of the admission time for the emergency. It references the *time* table.
departmentId	ID of the department of the hospital the emergency is assigned to. It references the *department* table.
genderId	ID of the gender of the patient. It references the *gender* table.
diagnosisId	ID of the diagnosis for the emergency. It references the *icd* table.
severityId	ID of the severity for the emergency. It references the *severity* table.
consultationCauseId	ID of the cause the patient goes to emergency for. It references the *consultationCause* table.

- *centre: the medical centres are stored in this table. It has the following columns:*

column	description
id	ID of the centre. Use this ID for referencing the centre from other tables.
management	Management unit for the centre, or '(Unknown)'.
area	Area where the centre is, or '(Unknown)'.
name	Name of the centre, or '(Unknown)'.

- *postalCode: the postal codes of the patients are stored in this table. It has the following columns:*

column	description
id	ID of the postal code. Use this ID for referencing the centre from other tables.
postalCode	The postal code itself, or '(Unknown)'.
lat	Latitude of the postal code location (used for GIS analyses), or NULL for '(Unknown)'.
lng	Longitude of the postal code location (used for GIS analyses), or NULL for '(Unknown)'.

- *dayOfWeek: the days of the week are stored in this table. It has the following columns:*

column	description
id	ID of the day of week. Use this ID for referencing the centre from other tables.
dayOfWeek	Name for the day of week in your language (Friday, vendredi, viernes...). It may be '(Unknown)', too
number	Ordinal for the day in the week. Use 1 for Sunday, 2 for Monday... 7 for Saturday, NULL for '(Unknown)'.

- *age: the ages of the patients are stored in this table. It has the following columns:*

column	description
id	ID of the age. Use this ID for referencing the centre from other tables.
ageGroup	Group this age belongs to. You can use these groups: 'Less than 5', '5 - 14', '15 - 29', '30 - 44', '45 - 59', '60 - 74', '75 - 89', '90 or more', '(Unknown)'.
ageGroupOrdinal	Ordinal for the age group. Use 0 for 'Less than 5', 1 for '5 - 14'... 7 for '90 or more', NULL for '(Unknown)'.
age	The age itself, or '(Unknown)'.
ageOrdinal	Ordinal for the age. Use the age number, except NULL for '(Unknown)'.

© SIDARTHa 2010

- *date: the dates of the emergencies are stored in this table. It has the following columns:*

column	description
id	ID of the date. Use this ID for referencing the centre from other tables.
year	The year, or '(Unknown)'.
quarter	The quarter. In the current version, it must be one the following values: 'Trimestre 1', 'Trimestre 2', 'Trimestre 3', 'Trimestre 4' or '(Unknown)'. We are working to improve this so you can see the quarters in your own language.
month	The month. In the current version, it must have the format 'mm - nnnnnn', where *mm* is the month number using two digits, and *nnnnnn* is the month name in the default locale of the Java VM. It may be '(Unknown)', too.
day_of_month	The day in the month, or '(Unknown)'.
date	The date itself, in MySQL native DATE format, or NULL for '(Unknown)'.

- *time: the times of the emergencies are stored in this table. It has the following columns:*

column	description
id	ID of the time. Use this ID for referencing the centre from other tables.
hour	The hour. It must have the format 'hh', where *hh* is the hour using two digits, from '00' to '23'. It may be '(Unknown)', too.
time	The time. It must have the format 'hh:mm', where *hh* is the hour using two digits, from '00' to '23', and *mm* is the minutes using two digits, from '00' to '59'. It may be '(Unknown)', too.
nativeTime	The time itself, in MySQL native TIME format, or NULL for '(Unknown)'.

- *department: the departments of the hospital are stored in this table. It has the following columns:*

column	description
Id	ID of the department. Use this ID for referencing the centre from other tables.
parentDepartment	Parent department this department belongs to, or '(Unknown)'.
department	The department itself, or '(Unknown)'.

- *gender: the times of the emergencies are stored in this table. It has the following columns:*

column	description
Id	ID of the gender. Use this ID for referencing the centre from other tables.
gender	Use 'FEMALE' for ID 1, 'MALE' for ID 2, '(Unknown)' for ID 3.

- *icd: the codes according to the ICD specification are stored in this table. We are currently using ICD-9-CM, as well as ICD-O-3. It has the following columns:*

column	description
Id	ID of the ICD. Use this ID for referencing the centre from other tables.
area	Area of the specification this code belongs to: 'MORPHOLOGY', 'PROCEDURAL' or 'DIAGNOSTICS'. It may be '(Unknown)', too.
level1	First level for grouping related codes, or '(Unknown)'.
level2	Second level for grouping related codes, or '(Unknown)'.
level3	Third level, that stores the code itself, or '(Unknown)'.

- *severity: the severities of the emergencies are stored in this table. It has the following columns:*

column	description
Id	ID of the severity. Use this ID for referencing the centre from other tables.
description	You can use the following values: 'SERIOUS', 'MODERATE', 'MILD' or '(Unknown)'.

- *consultationCause: the causes that the patients go to emergency for are stored in this table. It has the following columns:*

column	description
Id	ID of the cause. Use this ID for referencing the centre from other tables.
description	Description of the symptoms the patient has, or '(Unknown)'.

© SIDARTHa 2010